Identification of the Top 10 Core Genes as Diagnostic and Therapeutic Targets in Hepatocellular Carcinoma (HCC)

Qiangsheng Wu1, Yumei Yin2, Hui Gao1, Baoding Chen3, Jianjun Yu2, Jian Zhang4, Donglai Zhu5, Fang Liu6, Guohong Ge7 and Ya-Mei Dang8*

1Department of Assay Development, Eoto-Biotech Co., Ltd, Shenzhen, Guangdong 518000, China.
2Department of Medical Ultrasound, The Third People’s Hospital of Zhenjiang Affiliated Jiangsu University, Zhenjiang, Jiangsu 212005, China.
3Department of Medical Ultrasound, The Affiliated Hospital of Jiangsu University, Zhenjiang, Jiangsu 212001, China.
4Department of Radiology, The Affiliated Hospital of Jiangsu University, Zhenjiang, Jiangsu 212001, China.
5Department of Clinical Laboratory, The Third People’s Hospital of Zhenjiang Affiliated Jiangsu University, Zhenjiang, Jiangsu 212005, China.
6Department of Computer, Anhui Medical University, Hefei, Anhui 230032, China.
7Department of Hepatosis, The Third People’s Hospital of Zhenjiang Affiliated Jiangsu University, Zhenjiang, Jiangsu 212005, China.
8Department of Pathology, Gansu Provincial Hospital, Lanzhou, Gansu 730000, China.

Qiangsheng Wu, Yumei Yin and Hui Gao contributed equally to this work.

ABSTRACT

Lacking specific diagnostic markers for hepatocellular carcinoma is the critical reason for its increasing mortality. In this study, a microarray containing 24 hepatocellular carcinoma patients’ tumor tissues were compared to 8 normal liver tissues to screen out the significant abnormal expressed genes. The potential interactions and mechanisms of these genes in promoting hepatocellular carcinoma were analyzed through bioinformatics analysis. The results showed that the functions of these 1495 differently expressed genes enriched in cell cycle, oxidation-reduction and drug metabolic process. Additionally, the top 10 core genes of these differently expressed genes were mainly participated in cell cycle and strongly correlated with worse prognosis of survival, which may be potential diagnostic and therapeutic targets of HCC.

INTRODUCTION

Hepatocellular carcinoma (HCC) is a lethal malignancy of the liver. Primary liver cancer (85% was HCC) was the third leading cause of mortal cancer in the worldwide. Notably, 49.3% of the new cases and 58.3% of the deaths were occurred in Asia (Sung et al., 2021). HCC is associated with potentially alterable risk factors, such as excess body weight, alcohol abuse, hepatitis B virus (HBV), hepatitis C virus, nonalcohol fatty liver disease (Islami et al., 2018). As a high malignancy with insidious onset, invasive fast growing, high recurrence rate and fatality, HCC is a major public health problem especially in developing countries (McGlynn et al., 2021).

Effective and sensitive diagnostic and curative interventions available for HCC can serve to reduce the future burden and suffering of patients (Yang et al., 2019). Currently, diagnosis available for HCC screening relays on radio graphic tests and serological markers. However, most small nodules less than 2cm are extremely insensitive to be imaging detected, as well as the sensitivity of the serum golden marker AFP only ranges of 21%-64%
(even lower with small lesions) (Yang and Heimbach, 2020). The unfortunate truth is that it is often too late stages when patients have become symptomatic and been diagnosed, while no effective treatment would improve survival virtually (Singal et al., 2020; Yang et al., 2019). There is absolutely a need for higher sensitive and specific diagnostic markers for HCC.

Recently, bioinformatics analysis based on microarrays and high throughput sequencing technologies has been widely used to identifying differentially expressed genes (DEGs) and pathways related to the occurrence and developments of various cancers (Zhang et al., 2021). In this article, we provide comprehensive biological analysis of new potential targets of HCC and their functions in promoting and regulating tumor progression. Which hopefully will have unique contributions to ensure optimal treatments and outcomes for patients with HCC.

MATERIALS AND METHODS

Microarray and differently expressed genes analysis

The original data of GSE101685 microarray was downloaded from GEO datasets (https://www.ncbi.nlm.nih.gov/geo/). The gene expression profiles of 8 normal liver and 24 hepatocellular carcinoma tissues were selected and compared. Genes with adjusted $|\log 2 \text{FC}| \geq 1$ and $p$ value $< 0.05$ were identified as differentially expressed genes (DEGs).

Functional enrichment analysis

The differentially expressed gene list was uploaded to databases of DAVID 6.8 (https://david.ncifcrf.gov/home.jsp) and Metascape (https://metascape.org) separately (Zhou et al., 2019a). These two databases provided a comprehensive set of functional annotation for the biological meaning, which contained the Gene Ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis. In addition, the GO enrichment analysis included biological processes (BP), cellular components (CC), and molecular function (MF).

Neighbor genes and interaction network analysis

The differentially expressed gene list was uploaded to an open source software platform (Cytoscape) (Warde-Farley et al., 2010). Plugin STRING was used for analyzing the interaction network of DEGs. Plugin mCODE was used for finding the most highly interconnected region in DEGs. Plugin cytoHubba was used for exploring the most important nodes/hubs.

Top 10 core genes analysis (Shannon et al., 2003)

Interaction network of neighbor genes associated with the top 10 core genes was revealed on GeneMANIA (http://genemania.org/). Genetic alteration analysis of the top 10 core genes in HCC was explored from cbioPortal (www.cbioportal.org) based on TCGA database.

Aberrant expression analysis

Transcriptional levels of the top 10 core genes in 371 HCC and 50 normal tissues were explored and compared using UALCAN. Student’s $t$ test was used to generate a $p$ value. ****, $p<0.0001$.

Prognostic value analysis (Tang et al., 2017)

Correlative prognostic analysis of the top 10 core genes expression was performed using a Kaplan-Meier curve on GEPIA (http://gepia.cancer-pku.cn/index.html). Disease-free survival and overall survival curves were analyzed separately.

RESULTS

Gene expression profile of normal and tumor tissues of HCC

The microarray GSE101685 containing 8 normal liver tissues and 24 HCC patients’ tumor tissues was selected and showed the gene expression profile (Fig. 1A). 1495 different expressed genes were obtained with both the adjusted $|\log 2 \text{FC}| \geq 1$ and $p$ value $< 0.05$. Compared with the normal tissue, these genes different-expressed in HCC tissue include 650 up-regulated genes and 845 down-regulated genes (Fig. 1B).

Fig. 1. Different-expressed genes in HCC tumor tissue. (A) The gene expression profile in normal liver tissue and HCC tissue. (B) Statistically significant up-regulated or down-regulated expressed genes (red) in HCC tissue.

Functional enrichment analysis of DEGs in HCC

DAVID 6.8 and Metascape were utilized to investigate the functions of these differentially expressed genes. Figure 2A, B showed the most highly enriched GO and KEGG items of the significant up-regulated expressed
genes. Among the most highly enriched functions in the BP, CC and MF category or in KEGG pathways. Cell cycle was the most central functional item of all the up-regulated expressed genes. Figure 2C, D showed the most highly enriched GO and KEGG items of the significant down-regulated expressed genes. All these genes were mainly involved in oxidation-reduction and drug metabolic process.

Fig. 2. Enrichment analysis of the DEG in HCC. Bar plot of GO (BP, CC, MF) and KEGG enriched items of up-regulated genes analyzed by DAVID 6.8 (A) and Metascape (B). Bar plot of GO (BP, CC, MF) and KEGG enriched items of down-regulated genes analyzed by DAVID 6.8 (C) and Metascape (D).

The interaction network, mCODE and the top 10 core genes of the DEGs

To better understand the correction between these DEGs and HCC, the PPI network and the most significant mCODE component were analyzed (Fig. 3A, B). Then the Cytoscape plugin cytoHubba further ranked and showed the top 10 core genes. These core genes included NCAPG, NUF2, CCNB1, KIF11, BUB1B, RRM2, TTK, CCNB2, CCNA2 and MELK (Fig. 3C).

Neighbor gene network, functional enrichment and genetic alteration of the top 10 core genes

We further performed a comprehensive analysis of these 10 core genes. Firstly, Gene MANIA showed the neighbor gene network and revealed the functions of these core genes were primarily related to mitotic nuclear division, chromosome segregation and cell cycle checkpoint (Fig. 4A). While the functional enrichment was checked by Cytoscape. As expected, the core genes focused on cell cycle (Fig. 4B). Subsequently, provisional datasets of TCGA were utilized to analyze the genetic alterations. NCAPG, NUF2, CCNB1, KIF11, BUB1B, RRM2, TTK, CCNB2, CCNA2 and MELK were altered in 8%, 27%, 9%, 6%, 7%, 9%, 8%, 9%, 4% and 8% of the queried HCC samples. And mRNA high expression was the most common change in these samples (Fig. 4C).

Fig. 3. Interaction network, mCODE and the top 10 core genes analyses of the DEGs in HCC. (A) Protein-protein interaction network of the DEGs in HCC (STRING). (B, C) The most significant mCODE and top 10 core genes of the DEGs in HCC (cytoscape).

Fig. 4. Neighbor gene network, functional enrichment and genetic alteration of the top 10 core genes in HCC. Gene-gene interaction network (A) and the functional enrichment (B) of the top 10 core genes and their neighbor genes. (C) Summary of genetic alteration of the top 10 core genes in HCC.
Aberrant expression of the top 10 core genes in patients with HCC

To assess the expression levels of these top 10 core genes in larger sample size of HCC, the transcriptional levels of these genes in 371 HCC and 50 normal tissues were explored with UALCAN based on TCGA database (Fig. 5). As expected, the transcriptional levels of NCAPG, NUF2, CCNB1, KIF11, BUB1B, RRM2, TTK, CCNB2, CCNA2 and MELK in HCC tissues were significantly elevated (****, \( p < 0.0001 \)). Which indicated these genes have potential applications in the field of HCC diagnosis.

Fig. 5. The transcriptional levels of the top 10 core genes in HCC. Compared the transcription of NCAPG, NUF2, CCNB1, KIF11, BUB1B, RRM2, TTK, CCNB2, CCNA2 and MELK of HCC tissues with the normal tissues (UALCAn). LIHC represents liver hepatocellular carcinoma. ****, \( p < 0.0001 \).

The prognostic value of the top 10 core genes in patients with HCC

To evaluate the value of differentially expressed core genes in the progression of HCC, we assessed the correlation between the top 10 core genes and clinical outcome using GEPIA. Disease-free survival and overall survival curves are presented in Figure 6. HCC patients with low transcriptional levels of NCAPG (\( p=0.0024 \)), NUF2 (\( p=0.0038 \)), CCNB1 (\( p=0.000028 \)), KIF11 (\( p=0.00024 \)), BUB1B (\( p=0.00072 \)), RRM2 (\( p=0.00032 \)), TTK (\( p=0.0088 \)), CCNB2 (\( p=0.0064 \)), CCNA2 (\( p=0.0037 \)) and MELK (\( p=0.0014 \)) were significantly associated with longer disease-free survival (Fig. 6A). HCC patients with low transcriptional levels of NCAPG (\( p=0.00097 \)), NUF2 (\( p=0.0002 \)), CCNB1 (\( p=0.00015 \)), KIF11 (\( p=0.00061 \)), BUB1B (\( p=0.0028 \)), RRM2 (\( p=0.00058 \)), TTK (\( p=0.0015 \)), CCNA2 (\( p=0.0037 \)) and MELK (\( p=0.0015 \)) were significantly associated with better overall survival, except the CCNB2 (Fig. 6B). All these data demonstrate the differently expressed top 10 core genes may play significant roles and be potential therapeutic targets in HCC.

DISCUSSION

Although amount of basic and clinical studies has been performed to reveal the underlying molecular mechanism of HCC, the HCC-related genes have yet to be identified to elucidate the correlated with the cancer susceptibility, progression, and prognosis. This is mainly because the heterogeneity of the sample in independent studies or the single cohort nature in most of the previous studies. In this study, a microarray data containing 8 normal liver tissues and 24 HCC patients’ tumor tissues was analyzed, and a total of 1495 DEGs (650 upregulated and 845 down regulated) were identified in HCC tissues compared with normal liver tissues. For the further investigation of these DEGs, GO function and KEGG pathway analysis of these DEGs were performed.

Fig. 6. The prognostic value of the top 10 core genes in patients with HCC. The disease-free survival (A) and overall survival (B) curves of NCAPG, NUF2, CCNB1, KIF11, BUB1B, RRM2, TTK, CCNB2, CCNA2 and MELK in HCC.

GO analysis showed that upregulated DEGs were mainly enriched in cellular component groups, including nucleoplasm, chromosome, nucleus, and down-regulated DEGs were involved in organelle membrane, extracellular region, and mitochondrial matrix. Moreover, the KEGG pathways of upregulated DEGs were involved in cell cycle, DNA replication, Fanconi anemia pathway, and p53 signaling pathway, while the downregulated DEGs included metabolic pathways, fatty acid degradation, tryptophan metabolism, and complement and coagulation cascades. Afterward, the DEG PPI network complex was constructed, ten DEGs with the highest degrees of interaction in the PPI network complex were filtered.

To better understand the correction between these
DEGs and HCC, the neighbor Gene Network, functional enrichment, and genetic alteration of the top ten core genes was analyzed. The results revealed that the functions of these core genes were primarily related to mitotic nuclear division, chromosome segregation and cell cycle checkpoint. Finally, the aberrant expression of the 10 core genes shown that the transcriptional levels of NCAPG, NUF2, CCNB1, KIF11, BUB1B, RRM2, TTK, CCNB2, CCNA2 and MELK were significantly elevated in HCC tissues. In addition, the patients with dysregulation of the ten core genes had a worse prognosis, except the CCNB2. Consulting the latest research on the top 10 core genes, we found these genes were significantly connected with multiple cancers.

NCAPG (non-SMC condensin I complex subunit G), a subunit of the condensin complex, is associated with transcription in mitotic chromosome condensation. The increasing evidence reported that aberrant expression of NCAPG in various tumors, the inhibition of NCAPG can reduce Wnt/β-catenin signaling to impair the growth of endometrial cancer cells (Liu et al., 2021). NCAPG-related stimulation could activate PI3K/Akt pathway in the development of cardia adenocarcinoma (Zhang et al., 2020a). NCAPG has been found to be overexpressed in HCC compared with para-carcinoma tissue and is closely associated with poor prognosis and tumor immune infiltration in HCC patients with vascular (Liu et al., 2017b). Recently study have reported that NCAPG was significantly associated with tumor cell proliferation and migration in HCC (Zhang et al., 2018). Silencing NCAPG inhibits proliferation and induces apoptosis in HCC cells (Liu et al., 2018). Nevertheless, the mechanism by which NCAPG promotes proliferation in HCC remains unknown. Interestingly, the overexpression of NCAPG is closely related to the overexpression of CCNB1, which is involved in the top ten core genes of HCC (Wang et al., 2019a).

NUF2, a component of NDC80 kinetochore complex is essential for kinetochore-microtubule attachment and chromosome segregation. The recent study demonstrated that high NUF2 transcript expression can be a good prognostic biomarker to predict early tumor recurrence post-surgical resection in HCC. And the inhibition of NUF2 reduced the proliferation, migration, and invasion of HCC cells in vitro (Wang et al., 2019b). NuF2 regulated cell apoptosis and proliferation by regulating the binding of centromere and spindle microtubules to achieve the correct separation of chromosomes (Hu et al., 2015). The inhibition of NUF2 has been reported to reduce tumor cell growth in cancers of the brain, colon, liver, and pancreas (Huang et al., 2014). However, the role of NuF2 in the development of HCC remains uncertain.

CCNB1/CCNB2/CCNA2 as members of cyclin family, act as recognized oncogenes in cervical cancer, breast cancer, colorectal cancer and pancreatic cancer (Sugimasa et al., 2015). CCNB1 (cyclin B1) and CCNB2 (cyclin B2) can form complexes with cyclin-dependent kinase 1 (CDK1) to regulate the G2/M phases of the mammalian cell cycle, which plays an important role in the initiation of mitosis (Malumbres and Barbacid, 2005).

As shown in recent studies, the mRNA expression levels of CCNB1 and CCNB2 were significantly higher in several types of cancer and were associated with poor prognosis (Ding et al., 2014). Interestingly, in our study HCC patients with low transcriptional levels of CCNB2 associated with longer disease-free survival but not associated with overall survival. Knockdown of CCNB1 or CCNB2 could inhibit cell proliferation, invasion, migration, and G2/M phase progression of HCC cell lines and induced their apoptosis (Fang et al., 2014). As a cyclin controlling the G1/S and G2/M phases in the cell cycle, CCNA2 has been reported to participate in HCC development (Yam et al., 2002). The recent study indicated that overexpressed level of CCNA2 with other cell cycle regulators correlated with poor survival condition HCC subclass (Bai et al., 2020), which was in agreement with our prognostic analysis results. Other studies shown that CCNA2 regulated cell cycle and promote antiapoptotic activity in HCC patients (Qian et al., 2015).

Kinesin family member 11 (KIF11, also known as EG5), belongs to the kinesin-like protein family, is mainly involved in various kinds of spindle dynamics KIF11 functions as an oncogene in breast cancer, glioblastoma and oral cancer (Yang et al., 2009). The recent research demonstrated that high KIF11 expression in HCC tissues, which is strongly associated with liver cirrhosis and tumor stages and can be considered as a biomarker for poor prognosis of HCC (Liu et al., 2017a). Additionally, KIF11 interact with ASPM could promote the malignant progression of HCC via the Wnt/β-catenin signaling pathway (Wu et al., 2021). PRC1 plays a carcinogenic role by regulating KIF11 and promoting the occurrence and development of HCC (Chen et al., 2016). The in vivo study demonstrated that overexpression of KIF11 causes genomic instability in tumor cells and prompts abnormal proliferation in mice (Castillo et al., 2007).

BUB1B (BUB1 mitotic checkpoint serine/threonine kinase B) encodes a kinase involved in spindle checkpoint function. As a significant checkpoint of cell mitosis, BUB1B is involved in process of cell mitosis (Shin et al., 2003), and upregulation of BUB1B is observed in various human malignancies (Ando et al., 2010). Growing evidence shows that aberrant expression of BUB1B was highly involved in the tumorigenesis and development various cancers, for instance, overexpression of BUB1B...
is related to gastric cancer (Hanks et al., 2004), colon cancer (Abal et al., 2007), brain tumor (Ding et al., 2013), glioblastoma (Lee et al., 2017) and breast cancer (Scintu et al., 2007). Meanwhile, low expression of BUB1B associated metastasis and poorer survival have been found in colon adenocarcinoma (Shichiri et al., 2002) and lung tumor (Park et al., 2007). Recent study demonstrated that activation of FoxM1/BUB1B signaling pathway was vital for rhabdomyosarcoma growth and survival (Wan et al., 2012). Nevertheless, we still know quite little about the functional roles and potential mechanisms of BUB1B in HCC.

Human ribonucleotide reductase (RR) is a heterotetrametric complex by two large subunits RRM1 (ribonucleotide reductase catalytic subunit M1) and two small subunits RRM2 (ribonucleotide reductase regulatory subunit M2) (Heidel et al., 2007). RR maintains the homeostasis of nucleotide pools by converting ribonucleoside diphosphate to 20 deoxyribonucleoside diphosphate (Nordlund and Reichard, 2006). An-cancer expression profiling studies demonstrated that the expression of RRM2 is upregulated in multi types of cancers (Ding et al., 2019; Shao et al., 2013), dramatically enhances cellular invasiveness, angiogenesis, and metastasis (Zhou et al., 1998). In HCC tissues the expression of RRM2 is upregulated compared with adjacent normal liver tissues (Wang et al., 2010). The recent study revealed that the upregulation of RRM2 is closely associated with the poor prognosis of HCC patients, and RRM2 is enriched in the p53 signaling pathway (Zhou et al., 2019b). Additionally, Several RRM2 inhibitors are currently clinically used for anticancer and antivirus therapy (Shao et al., 2013). RRM2 siRNA transfection efficiently inhibited the cell proliferation of HCC cells (Yang et al., 2020). The in vivo study shows that the RRM2 siRNA alone or in combination with adriamycin could inhibit xenografted or orthotopic tumor growth established in nude mice (Gao et al., 2013).

TTK encodes a dual specificity protein kinase associated with cell proliferation. Differently expressed TTK may result in aberrant mitotic spindles and multiple transcript variants (Mills et al., 1992). Accumulated evidence shown that overexpression of TTK could be detected in different types of cancer, including HCC (Jemal et al., 2011; Xie et al., 2017), glioblastoma (Xie et al., 2017), esophageal cancer (Mizukami et al., 2008), and breast cancer (Fan et al., 2006). In our study, TTK was overexpressed in HCC tissues on mRNA levels and HCC patients with high expression levels of TTK had lower survival rates. The recent research had speculated that TTK in HCC carcinogenesis by promoting cell survival and invasion via activation of Akt/mTOR and MDM2/ p53 signaling pathways, and the regulation of miR-21 via TGF-β (Liu et al., 2015). In addition, TTK plays an essential role in proliferation and sorafenib resistance of HCC cells and could act as a biomarker and potential target for HCC treatment (Liang et al., 2014).

Maternal embryonic leucine zipper kinase (MELK) is a unique member of the Snf1/AMPK kinase family. It belongs to a conservative cycle-dependent kinase (Li et al., 2022). MELK plays an important role in cell cycle regulation, cell proliferation and other processes (Nakano et al., 2005). MELK is highly expressed in a variety of human tumors, the recent study shown that E2F1/MELK/mTORC1/2 axis involved in the progression of endometrial carcinoma (EC) (Xu et al., 2020). The inhibition of MELK could suppress the growth of Glioblastoma multiforme (GBM) via blockage AKT signals (Zhang et al., 2020b). Numerous studies regarding the development of HCC provide evidence that variation in expression level of MELK is concerned with the development of HCC (Xia et al., 2016). In our study, the analysis of HCC patients identified that mRNA level of MELK was highly overexpressed, which was correlated with early recurrence and poor overall survival. Further study demonstrated that silencing MELK inhibited the cell growth, invasion, stenness and tumorigenicity of HCC cells by inducing apoptosis and mitosis (Xia et al., 2016).

Overall, by using a stepwise analysis of a microarray of 24 HCC tumor tissues and 8 normal liver tissues, we identified 1495 DEGs candidate genes (650 upregulated and 845 downregulated). Ten core genes with the highest interaction degrees were obtained, and these genes were significantly enriched in several pathways, mainly related to mitotic nuclear division, chromosome segregation and cell cycle checkpoint. The aberrant expression and survival analysis also showed a close interaction between these core genes and the poor prognosis of HCC.

In this study, the top 10 core genes NCA PG, NUF 2, CCNB1, KIF11, BUB1B, RRM2, TTK, CCNB2, CCNA2 and MELK might be important oncogenes of cell cycle in HCC. These findings could significantly enrich our understanding of the development and recurrence of HCC. Moreover, these genes could be therapeutic targets for HCC treatment. Future studies will be needed to further understand the underlying mechanisms of the concrete interaction of these genes.

ACKNOWLEDGMENTS

This work was supported by grant from Science and Technological Development Foundation of Zhenjiang (SH2019086, SH2020035, FZ2020050).
Statement of conflict of interest

The authors have declared no conflict of interest.

REFERENCES


Issa, F., Goding Sauer, A., Miller, K., Siegel, R., Fedewa, S., Jacobs, E., McCullough, M., Patel, A., Ma, J., Soerjomataram, I., Flanders, W., Brawley,


Identification of the Top 10 Core Genes as Diagnostic and Therapeutic Targets


